

On the largest-eigenvalue process for generalized Wishart random matrices

A. B. Dieker and J. Warren

Georgia Institute of Technology, Atlanta GA 30332, USA

E-mail address: `ton.dieker@isye.gatech.edu`

University of Warwick, Department of Statistics, Coventry, CV4 7AL, United Kingdom

E-mail address: j.warren@warwick.ac.uk

Abstract. Using a change-of-measure argument, we prove an equality in law between the process of largest eigenvalues in a generalized Wishart random-matrix process and a last-passage percolation process. This equality in law was conjectured by Borodin and P  ch   (2008).

1. Introduction

The past decade has witnessed a surge of interest in connections between random matrices on the one hand and applications to growth models, queueing systems, and last-passage percolation models on the other hand; standard references are Baryshnikov (2001) and Johansson (2000). In this note we prove a result of this kind: an equality in law between a process of largest eigenvalues for a family of Wishart random matrices and a process of directed last-passage percolation times.

To formulate the main result, we construct two infinite arrays of random variables on an underlying measurable space, along with a family $\{P^{\pi, \hat{\pi}}\}$ of probability measures parametrized by a positive N -vector π and a nonnegative sequence $\{\hat{\pi}_n : n \geq 1\}$. The elements of the first array $\{A_{ij} : 1 \leq i \leq N, j \geq 1\}$ are independent and A_{ij} has a complex zero-mean Gaussian distribution with variance $1/(\pi_i + \hat{\pi}_j)$ under $P^{\pi, \hat{\pi}}$. That is, both the real and complex part of A_{ij} have zero mean and variance $1/(2\pi_i + 2\hat{\pi}_j)$. Write $A(n)$ for the $N \times n$ matrix formed by the first n columns of A , and define the matrix-valued stochastic process $\{M(n) : n \geq 0\}$ by setting $M(n) = A(n)A(n)^*$ for $n \geq 1$ and by letting $M(0)$ be the $N \times N$ zero matrix. We call $\{M(n) : n \geq 0\}$ a generalized Wishart random-matrix process, since the marginals have a Wishart distribution if π and $\hat{\pi}$ are identically one and zero, respectively.

The elements of the second array $\{W_{ij} : 1 \leq i \leq N, j \geq 1\}$ are independent and W_{ij} is exponentially distributed with parameter $\pi_i + \hat{\pi}_j$ under $P^{\pi, \hat{\pi}}$. We define

$$Y(N, n) = \max_{P \in \Pi(N, n)} \sum_{(ij) \in P} W_{ij},$$

The following theorem, a process-level equality in law between the largest eigenvalue of $M(n)$ and $Y(N, n)$, is the main result of this note. Given a matrix C , we write $\text{sp}(C)$ for its vector of eigenvalues, ordered decreasingly.

It is known from Defosseux (2008); Forrester and Rains (2006) that this holds in the ‘standard’ case, i.e., under the measure $P := P^{(1,\dots,1),(0,0,\dots)}$. In its stated generality, the theorem was conjectured by Borodin and Péché (2008), who prove that the laws of $Y(N, n)$ and the largest eigenvalue of $M(n)$ coincide for fixed $n \geq 1$. Our proof is based on a change-of-measure argument, which is potentially useful to prove related equalities in law.

$$x'_1 \geq x_1 \geq x'_2 \geq x_2 \geq \dots \geq x'_N \geq x_N.$$

This section provides some background on generalized Wishart random matrices, and introduces a Markov chain which plays an important role in the proof of Theorem 1.1.

$$M(m) = M(m-1) + (A_{im}\bar{A}_{jm})_{1 \leq i,j \leq N}, \quad (2.1)$$

Proposition 2.1. *For any $m \geq 1$, the $P^{\pi, \hat{\pi}}$ -law of $M(m) - M(m-1)$ is absolutely continuous with respect to the P -law of $M(m) - M(m-1)$, and the Radon-Nikodym derivative is*

$$\prod_{i=1}^N (\pi_i + \hat{\pi}_m) \exp \left(- \sum_{i=1}^N (\pi_i + \hat{\pi}_m - 1) (M_{ii}(m) - M_{ii}(m-1)) \right).$$

2.2. A Markov transition kernel. We next introduce a time-inhomogeneous Markov transition kernel on W^N . We shall prove in Section 3 that this kernel describes the eigenvalue-process of the generalized Wishart random-matrix process of the previous subsection.

In the standard case ($\pi \equiv 1$, $\hat{\pi} \equiv 0$), it follows from unitary invariance (see Defosseux (2008, Sec. 5) or Forrester and Rains (2006)) that the process $\{\text{sp}(M(n)) : n \geq 0\}$ is a homogeneous Markov chain. Its one-step transition kernel $Q(z, \cdot)$ is the law of $\text{sp}(\text{diag}(z) + G)$, where $G = \{g_i \bar{g}_j : 1 \leq i, j \leq N\}$ is a rank one matrix determined by an N -vector g of standard complex Gaussian random variables. For z in the interior of W^N , $Q(z, \cdot)$ is absolutely continuous with respect to Lebesgue measure on W^N and can be written explicitly as in Defosseux (2008, Prop. 4.8):

$$Q(z, dz') = \frac{\Delta(z')}{\Delta(z)} e^{-\sum_k (z'_k - z_k)} 1_{\{z \prec z'\}} dz',$$

where $\Delta(z) := \prod_{1 \leq i < j \leq N} (z_i - z_j)$ is the Vandermonde determinant.

We use the Markov kernel Q to define the aforementioned time-inhomogeneous Markov kernels, which arise from the generalized Wishart random-matrix process. For general π and $\hat{\pi}$, we define the inhomogeneous transition probabilities $Q_{n-1,n}^{\pi, \hat{\pi}}$ via

$$Q_{n-1,n}^{\pi, \hat{\pi}}(z, dz') = \prod_{i=1}^N (\pi_i + \hat{\pi}_n) \frac{h_\pi(z')}{h_\pi(z)} e^{-(\hat{\pi}_n - 1) \sum_{i=1}^N (z'_i - z_i)} Q(z, dz'),$$

with

$$h_\pi(z) = \frac{\det\{e^{-\pi_i z_j}\}}{\Delta(\pi)\Delta(z)}. \quad (2.2)$$

Note that $h_\pi(z)$ extends to a continuous function on $(0, \infty)^N \times W^N$ (this can immediately be seen as a consequence of the Harish-Chandra-Itzykson-Zuber formula, see (3.2) below).

One can verify that the $Q^{\pi, \hat{\pi}}$ are true Markov kernels by writing $1_{\{z \prec z'\}} = \det\{1_{\{z_i < z'_j\}}\}$ and applying the Cauchy-Binet formula

$$\int_{W^N} \det\{\xi_i(z_j)\} \det\{\psi_j(z_i)\} dz = \det\left\{\int_{\mathbb{R}} \xi_i(z) \psi_j(z) dz\right\}.$$

3. The generalized Wishart eigenvalue-process

In this section, we determine the law of the eigenvalue-process of generalized Wishart random-matrix process. Although it is not essential to the proof of Theorem 1.1, we formulate our results in a setting where $\text{sp}(M(0))$ is allowed to be nonzero.

Write m_μ for the ‘uniform distribution’ on the set $\{M \in \mathbf{H}_{N,N} : \text{sp}(M) = \mu\}$. That is, m_μ is the unique probability measure invariant under conjugation by unitary matrices, or equivalently m_μ is the law of $U \text{diag}(\mu) U^*$ where U is unitary and distributed according to (normalized) Haar measure. We define measures $P_\mu^{\pi, \hat{\pi}}$ by letting the $P_\mu^{\pi, \hat{\pi}}$ -law of $\{M(n) - M(0) : n \geq 0\}$ be equal to the $P^{\pi, \hat{\pi}}$ -law of $\{M(n) : n \geq 0\}$, and letting the $P_\mu^{\pi, \hat{\pi}}$ -distribution of $M(0)$ be independent of $\{M(n) - M(0) : n \geq 0\}$ and absolutely continuous with respect to m_μ with Radon-Nikodym derivative

$$\frac{c_N}{h_\pi(\mu)} e^{-\sum_{i=1}^N \mu_i} \exp(-\text{tr}[(\text{diag}(\pi) - I)M(0)]), \quad (3.1)$$

where c_N is a constant depending only on the dimension N and I is the identity matrix. Recall that $h_\pi(\mu)$ is defined in (2.2). That this defines the density of a probability measure for all π and μ follows immediately from the Harish-Chandra-Itzykson-Zuber formula (e.g., Mehta (2004, App. A.5))

$$\int_U \exp(-\text{tr}(\text{diag}(\pi)U \text{diag}(\mu)U^*)) dU = c_N^{-1} h_\pi(\mu), \quad (3.2)$$

writing dU for normalized Haar measure on the unitary group. Throughout, we abbreviate $P_\mu^{(1,\dots,1),(0,0,\dots)}$ by P_μ . Note that the $P_\mu^{\pi,\hat{\pi}}$ -law and the $P^{\pi,\hat{\pi}}$ -law of $\{M(n) : n \geq 0\}$ coincide if $\mu = 0$.

The following theorem specifies the $P_\mu^{\pi,\hat{\pi}}$ -law of $\{\text{sp}(M(n)) : n \geq 0\}$.

Theorem 3.1. *For any $\mu \in W^N$, $\{\text{sp}(M(n)) : n \geq 0\}$ is an inhomogeneous Markov chain on W^N under $P_\mu^{\pi,\hat{\pi}}$, and it has the $Q_{n-1,n}^{\pi,\hat{\pi}}$ of Section 2.2 for its one-step transition kernels.*

Proof. Fix some $\mu \in W^N$. The key ingredient in the proof is a change of measure argument. We know from Defosseux (2008) or Forrester and Rains (2006) that Theorem 3.1 holds for the ‘standard’ case $\pi = (1, \dots, 1)$, $\hat{\pi} \equiv 0$.

Writing $P_n^{\pi,\hat{\pi}}$ and P_n for the distribution of $(M(0), \dots, M(n))$ under $P_\mu^{\pi,\hat{\pi}}$ and P_μ respectively, we obtain from Section 2.1 that for $n \geq 0$,

$$\begin{aligned} & \frac{dP_n^{\pi,\hat{\pi}}}{dP_n}(M(0), \dots, M(n)) \\ &= C_{\pi,\hat{\pi}}(n, N) \frac{c_N}{h_\pi(\mu)} e^{-\sum_{i=1}^N \mu_i} \\ & \quad \times \exp \left(-\text{tr}((\text{diag}(\pi) - I)M(n)) - \sum_{m=1}^n \hat{\pi}_m \text{tr}(M(m) - M(m-1)) \right), \end{aligned}$$

where $C_{\pi,\hat{\pi}}(n, N) = \prod_{i=1}^N \prod_{j=1}^n (\pi_i + \hat{\pi}_j)$. Let the measure $p_n^{\pi,\hat{\pi}}$ (and p_n) be the restriction of $P_n^{\pi,\hat{\pi}}$ (and P_n) to the σ -field generated by $(\text{sp}(M(0)), \dots, \text{sp}(M(n)))$. Then we obtain for $n \geq 0$,

$$\begin{aligned} & \frac{dp_n^{\pi,\hat{\pi}}}{dp_n}(\text{sp}(M(0)), \dots, \text{sp}(M(n))) \\ &= \mathbb{E}_{P_\mu} \left[\frac{dP_n^{\pi,\hat{\pi}}}{dP_n}(M(0), \dots, M(n)) \middle| \text{sp}(M(0)), \dots, \text{sp}(M(n)) \right], \end{aligned}$$

where \mathbb{E}_{P_μ} denotes the expectation operator with respect to P_μ . Since the P_μ -distribution of $(M(0), \dots, M(n))$ given the spectra is invariant under component-wise conjugation by a unitary matrix U , we have for $\mu \equiv \mu^{(0)} \prec \mu^{(1)} \prec \dots \prec \mu^{(n)}$,

$$\begin{aligned} & \mathbb{E}_{P_\mu} \left[\exp(-\text{tr}(\text{diag}(\pi)M(n))) \middle| \text{sp}(M(0)) = \mu^{(0)}, \dots, \text{sp}(M(n)) = \mu^{(n)} \right] \\ &= \int_U \exp \left(-\text{tr}(\text{diag}(\pi)U \text{diag}(\mu^{(n)})U^*) \right) dU \\ &= c_N^{-1} h_\pi(\mu^{(n)}), \end{aligned}$$

where the second equality is the Harish-Chandra-Itzykson-Zuber formula. From the preceding three displays in conjunction with $\text{tr}(M) = \sum_i \text{sp}(M)_i$, we conclude

that

$$\begin{aligned} & \frac{dp_n^{\pi, \hat{\pi}}}{dp_n}(\mu, \mu^{(1)}, \dots, \mu^{(n)}) \\ &= C_{\pi, \hat{\pi}}(n, N) \frac{h_{\pi}(\mu^{(n)})}{h_{\pi}(\mu)} \exp \left(- \sum_{i=1}^N \sum_{r=1}^n \hat{\pi}_r [\mu_i^{(r)} - \mu_i^{(r-1)}] + \sum_{i=1}^N [\mu_i^{(n)} - \mu_i] \right). \end{aligned}$$

Since $\text{sp}(M(\cdot))$ is a Markov chain with transition kernel Q under P_{μ} , we have

$$\begin{aligned} & P_{\mu}^{\pi, \hat{\pi}}(\text{sp}(M(1)) \in d\mu^{(1)}, \dots, \text{sp}(M(n)) \in d\mu^{(n)}) \\ &= \frac{dp_n^{\pi, \hat{\pi}}}{dp_n}(\mu, \mu^{(1)}, \dots, \mu^{(n)}) P_{\mu}(\text{sp}(M(1)) \in d\mu^{(1)}, \dots, \text{sp}(M(n)) \in d\mu^{(n)}) \\ &= \frac{dp_n^{\pi, \hat{\pi}}}{dp_n}(\mu, \mu^{(1)}, \dots, \mu^{(n)}) Q(\mu, d\mu^{(1)}) \dots Q(\mu^{(n-1)}, d\mu^{(n)}) \\ &= Q_{0,1}^{\pi, \hat{\pi}}(\mu, d\mu^{(1)}) Q_{1,2}^{\pi, \hat{\pi}}(\mu^{(1)}, d\mu^{(2)}) \dots Q_{n-1,n}^{\pi, \hat{\pi}}(\mu^{(n-1)}, d\mu^{(n)}), \end{aligned}$$

the last equality being a consequence of the definition of $Q_{k-1,k}^{\pi, \hat{\pi}}$ and the expression for $dp_n^{\pi, \hat{\pi}}/dp_n$. \square

4. Robinson-Schensted-Knuth and the proof of Theorem 1.1

This section explains the connection between the infinite array $\{W_{ij}\}$ of the introduction and the Markov kernels $Q_{n-1,n}^{\pi, \hat{\pi}}$. In conjunction with Theorem 3.1, these connections allow us to prove Theorem 1.1.

The RSK algorithm. The results in this section rely on a combinatorial mechanism known as the Robinson-Schensted-Knuth (RSK) algorithm. This algorithm generates from a $p \times q$ matrix with nonnegative entries a triangular array $\mathbf{x} = \{x_i^j : 1 \leq j \leq p, 1 \leq i \leq j\}$ called a Gelfand-Tsetlin (GT) pattern. A GT pattern with p levels x^1, \dots, x^p is an array for which the coordinates satisfy the inequalities

$$x_k^k \leq x_{k-1}^{k-1} \leq x_{k-1}^k \leq x_{k-2}^{k-1} \leq \dots \leq x_2^k \leq x_1^{k-1} \leq x_1^k$$

for $k = 2, \dots, p$. If the elements of the matrix are integers, then a GT pattern can be identified with a so-called semistandard Young tableau, and the bottom row $x^p = \{x_i^p; 1 \leq i \leq p\}$ of the GT pattern corresponds to the shape of the Young tableau. We write \mathbf{K}_p for the space of all GT patterns \mathbf{x} with p levels.

By applying the RSK algorithm with row insertion to an infinite array $\{\xi_{ij} : 1 \leq i \leq N, 1 \leq j \leq n\}$ for $n = 1, 2, \dots$, we obtain a sequence of GT patterns $\mathbf{x}(1), \mathbf{x}(2), \dots$. It follows from properties of RSK that

$$x_1^N(n) = \max_{P \in \Pi(N,n)} \sum_{(ij) \in P} \xi_{ij}, \quad (4.1)$$

where $\Pi(N, n)$ is the set of up-right paths from $(1, 1)$ to (N, n) as before. Details can be found in, e.g., Johansson (2000) or Dieker and Warren (2008, case A).

Greene's theorem generalizes (4.1), and gives similar expressions for each component of the pattern $x_i^j(n)$, see for instance Chapter 3 of Fulton (1997) or Equation (16) in Doumerc (2003). As a consequence of these, we can consider the RSK algorithm for real-valued ξ_{ij} and each $\mathbf{x}(n)$ is then a continuous function of the input data.

We remark that the RSK algorithm can also be started from a given initial GT pattern $\mathbf{x}(0)$. If RSK is started from the null pattern, it reduces to the standard algorithm and we set $\mathbf{x}(0) = 0$.

The bijective property of RSK. RSK has a bijective property which has important probabilistic consequences for the sequence of GT patterns constructed from specially chosen random infinite arrays. Indeed, suppose that $\{\xi_{ij} : 1 \leq i \leq N, j \geq 1\}$ is a family of independent random variables with ξ_{ij} having a geometric distribution on \mathbb{Z}_+ with parameter $a_i b_j$, where $\{a_i : 1 \leq i \leq N\}$ and $\{b_j : j \geq 1\}$ are two sequences taking values in $(0, 1]$. Write $\{\mathbf{X}(n) : n \geq 0\}$ for the sequence of GT patterns constructed from ξ .

Using the bijective property of RSK it can be verified that the bottom rows $\{X^N(n) : n \geq 0\}$ of the GT patterns evolve as an inhomogeneous Markov chain with transition probabilities

$$P_{n-1,n}(x, x') = \prod_{i=1}^N (1 - a_i b_n) \frac{s_{x'}(a)}{s_x(a)} b_n^{\sum_{i=1}^N (x'_i - x_i)} 1_{\{0 \leq x \prec x'\}}, \quad (4.2)$$

where $s_\lambda(a)$ is the Schur polynomial corresponding to a partition λ :

$$s_\lambda(a) = \sum_{\mathbf{x} \in \mathbf{K}_N : x^N = \lambda} a^{\mathbf{x}},$$

with the weight $a^{\mathbf{x}}$ of a GT pattern \mathbf{x} being defined as

$$a^{\mathbf{x}} = a_1^{x_1^1} \prod_{k=2}^N a_k^{\sum x_i^k - \sum x_i^{k-1}}.$$

This is proved in O'Connell (2003) in the special case with $b_j = 1$ for all j , and the argument extends straightforwardly; see also Forrester and Nagao (2008).

Non-null initial GT patterns generally do *not* give rise to Markovian bottom-row processes. Still, the inhomogeneous Markov chain of bottom rows can be constructed starting from a given initial partition λ with at most N parts by choosing $\mathbf{X}(0)$ suitably from the space of a GT patterns with bottom row λ : $\mathbf{X}(0)$ should be independent of the family $\{\xi_{ij}\}$ with probability mass function

$$p(\mathbf{x}) = \frac{a^{\mathbf{x}}}{s_\lambda(a)}.$$

Exponentially distributed input data. We now consider the sequence of GT patterns $\{\mathbf{X}_L(n) : n \geq 0\}$ arising from setting $a_i = 1 - \pi_i/L$ and $b_j = 1 - \hat{\pi}_j/L$ in the above setup, and we study the regime $L \rightarrow \infty$ after rescaling suitably. In the regime $L \rightarrow \infty$, the input variables $\{\xi_{ij}/L\}$ (jointly) converge in distribution to independent exponential random variables, the variable corresponding to ξ_{ij}/L having parameter $\pi_i + \hat{\pi}_j$. Thus, the law of the input array ξ converges weakly to the $P^{\pi, \hat{\pi}}$ -law of the array $\{W_{ij} : 1 \leq i \leq N, j \geq 1\}$ from the introduction. Refer to Doumerc (2003) and Johansson (2000) for related results on this regime.

By the aforementioned continuity of the RSK algorithm and the continuous-mapping theorem, $\{\mathbf{X}_L(n)/L : n \geq 0\}$ converges in distribution to a process $\{\mathbf{Z}(n) :$

$n \geq 0\}$ taking values in GT patterns with N levels. As a consequence of the above results in a discrete-space setting, we get from (4.1) that

$$Z_1^N(n) = \max_{P \in \Pi(N,n)} \sum_{(ij) \in P} W_{ij}.$$

Moreover, the process of bottom rows $\{Z^N(n) : n \geq 0\}$ is an inhomogeneous Markov chain for which its transition mechanism can be found by letting $L \rightarrow \infty$ in (4.2):

Lemma 4.1. *Under $P^{\pi, \hat{\pi}}$, the process $\{Z^N(n) : n \geq 0\}$ is an inhomogeneous Markov chain on W^N , and it has the $Q_{n-1,n}^{\pi, \hat{\pi}}$ of Section 2.2 for its one-step transition kernels.*

A similar result can be obtained given a non-null initial bottom row $\mu \in W^N$. In case the components of μ are distinct, the distribution of the initial pattern $\mathbf{Z}(0)$ should then be absolutely continuous with respect to Lebesgue measure on $\{\mathbf{z} \in \mathbf{K}_N : z^N = \mu\}$ with density

$$\frac{\Delta(\pi)}{\det\{e^{-\pi_i \mu_j}\}} c^{\mathbf{z}},$$

where $c = (e^{-\pi_1}, \dots, e^{-\pi_N})$.

Proof of Theorem 1.1. We now have all ingredients to prove Theorem 1.1. We already noted that $Z_1^N(n)$ equals $Y(N, n)$. Thus, for any strictly positive vector π and any nonnegative sequence $\hat{\pi}$, $\{Y(N, n) : n \geq 1\}$ has the same $P^{\pi, \hat{\pi}}$ -distribution as $\{Z_1(n) : n \geq 1\}$. In view of Theorem 3.1 and Lemma 4.1, in turn this has the same $P^{\pi, \hat{\pi}}$ -distribution as the largest-eigenvalue process $\{\text{sp}(M(n))_1 : n \geq 1\}$. This proves Theorem 1.1.

References

- Yu. Baryshnikov. GUEs and queues. *Probab. Theory Related Fields* **119**, 256–274 (2001).
- A. Borodin and S. Péché. Airy kernel with two sets of parameters in directed percolation and random matrix theory. *J. Stat. Phys.* **132**, 275–290 (2008).
- M. Defosseux. Orbit measures, random matrix theory and interlaced determinantal processes (2008). ArXiv:0810.1011.
- A. B. Dieker and J. Warren. Determinantal transition kernels for some interacting particles on the line. *Ann. Inst. H. Poincaré Probab. Statist.* **44**, 1162–1172 (2008).
- Y. Doumerc. A note on representations of eigenvalues of classical Gaussian matrices. In *Séminaire de Probabilités XXXVII*, volume 1832 of *Lecture Notes in Math.*, pages 370–384. Springer, Berlin (2003).
- P. J. Forrester and T. Nagao. Determinantal correlations for classical projection processes (2008). ArXiv:0801.0100.
- P. J. Forrester and E. M. Rains. Jacobians and rank 1 perturbations relating to unitary Hessenberg matrices. *Int. Math. Res. Not.* page Art. ID 48306 (2006).
- W. Fulton. *Young tableaux*. Cambridge University Press, Cambridge (1997).
- K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.* **209**, 437–476 (2000).
- K. Johansson. A multi-dimensional Markov chain and the Meixner ensemble. *Arkiv för Matematik, to appear* (2009). Doi:10.1007/s11512-008-0089-6.
- M. L. Mehta. *Random matrices*. Elsevier/Academic Press, Amsterdam, third edition (2004). ISBN 0-12-088409-7.
- N. O’Connell. Conditioned random walks and the RSK correspondence. *J. Phys. A* **36**, 3049–3066 (2003).